

Comparativa entre los resultados obtenidos mediante evaluación continua y evaluación final en materias técnicas online de posgrado. Influencia del factor tiempo y la puntuación de las actividades.

Roberto del Teso March^a, Elvira Estruch-Juan^a, Elena Gómez Sellés^a, Javier Soriano Olivares^a

^aITA, Departamento de Ingeniería Hidráulica y Medio Ambiente. Escuela Técnica Superior de Ingenieros industriales. Universitat Politècnica de València

Resumen

La metodología empleada para la evaluación de los procesos de aprendizaje es clave para determinar la adquisición de las competencias. En docencia online, la autoevaluación y el feedback emergen como elementos indispensables que, combinados con distintas actividades de aprendizaje, trabajan en línea para conseguir que el alumnado alcance los objetivos marcados. Este trabajo analiza los resultados conseguidos por los alumnos de posgrado de la formación online Cursosagua a lo largo de cuatro cursos académicos, para determinar cuáles son los factores más influyentes y mejorar el sistema de evaluación actual.

Palabras clave: *formación a distancia, evaluación online, formación técnica, metodología docente, seguimiento del aprendizaje*

1. Introducción

Una de las grandes dificultades a la hora de hacer efectiva la docencia superior a distancia en plataformas educativas, es el establecimiento de los sistemas de evaluación de la manera más adecuada (Rodríguez & Ibarra, 2011). Estos sistemas de evaluación deben ofrecer una calificación coherente con el aprendizaje de los estudiantes. Para ello, la autoevaluación en entornos virtuales es fundamental (García-Beltrán, Martínez, Jaén, & Tapia, 2006).

Con la autoevaluación, los estudiantes desarrollan la capacidad de juicio sobre su proceso de desarrollo (Gibbs & Simpson, 2009). Por este motivo, es importante ofrecer diferentes actividades autoevaluativas a lo largo del temario. El objetivo es que tanto el alumno como los tutores puedan hacer un seguimiento del proceso de aprendizaje, asegurando que se

alcanzan los resultados de aprendizaje estipulados, y en caso contrario, poder reorientar al alumno, y mejorar su proceso de aprendizaje. La retroalimentación que el estudiante recibe de cada una de las actividades planteada debe ofrecerse a tiempo.

En la docencia online, las plataformas educativas recogen información del proceso de aprendizaje de cada alumno (analíticas de aprendizaje). Esta información puede ser utilizada por los tutores como base para ofrecer feedback y orientación a los estudiantes a lo largo de su estudio (del Teso, Estruch, Gómez, & Soriano, 2018)(del Teso, Estruch, Gómez, & Soriano, 2018) Se deben diseñar diferentes actividades evaluables para que el feedback sea efectivo (Gibbs & Simpson, 2009).

En algunos cursos y titulaciones, las actividades autoevaluativas y el feedback recibido por el alumno, permiten consolidar conocimientos previos necesarios antes de introducir una nueva materia o temario. Este es el caso de la metodología utilizada por el ITA en su oferta formativa de posgrado online.

El ITA, grupo de investigación perteneciente al Departamento de Ingeniería Hidráulica y Medio Ambiente de la Universitat Politècnica de València, con su formación Cursosagua, cuenta con un Plan de Estudios Coordinado de docencia online en el que se imparten 23 asignaturas y 5 títulos de posgrado relacionados con la hidráulica y la gestión del agua. En estos títulos, la evaluación de cada una de las materias se divide en dos partes. Por un lado, una evaluación continua que representa el 50% de la nota final y que a su vez sirve de autoevaluación al alumno, pues se le ofrece de manera inmediata un feedback sobre su respuesta. Por otro lado, un examen final que representa el 50% restante de la nota.

Pese a que las actividades planteadas en las titulaciones del Plan de Estudios Coordinado son similares en la evaluación continua y en la evaluación final, se ha detectado una diferencia notoria entre las calificaciones obtenidas en ambas partes, siendo, en general, la nota del examen final considerablemente menor.

El objetivo de este trabajo es analizar estadísticamente las posibles causas de esta diferencia de notas, con el fin de mejorar la evaluación orientada al aprendizaje. Para ello, se cuenta con los resultados de 160 alumnos de posgrado recogidos durante 4 cursos académicos. Se analizará la diferencia de notas en función de la asignatura, el título que el alumno está realizando, el tiempo establecido para completar la asignatura, la convocatoria de examen en la que se ha realizado la asignatura y el tipo de actividades programadas. De este modo, será posible conocer qué factores tienen una influencia estadísticamente significativa sobre los resultados obtenidos y proponer mejoras en el proceso evaluativo.

2. El sistema de evaluación Cursosagua

Como se ha comentado previamente, la evaluación de cada uno de los módulos pertenecientes a los títulos de posgrado a distancia, se divide en dos partes: una evaluación continua que representa el 50% de la nota final y un examen final que representa el 50% restante.

La evaluación continua, a su vez, cuenta con diferentes actividades evaluables a lo largo del temario. Se busca que las preguntas del examen final sean coherentes con la tipología de actividades planteadas en la evaluación continua.

Las actividades que los estudiantes deben realizar en la parte de evaluación continua se dividen en cuatro tareas diferentes con su correspondiente puntuación:

Puntos de control: Son autoevaluaciones que se desarrollan a lo largo de cada unidad. Para que el estudiante pueda continuar con el temario, debe finalizar por completo el punto de control. Las preguntas realizadas en estos puntos de control tienen como objetivo mantener activo al estudiante durante su estudio (González de la Fuente y Carabantes Alarcón, 2018) y consolidar los conocimientos que ha adquirido hasta ese momento, recibiendo un feedback instantáneo tras cada respuesta.

Test de las unidades: Al finalizar cada una de las unidades que componen una asignatura, se habilita automáticamente un test con preguntas del tipo multirrespuesta sobre conceptos teóricos de la unidad. Este tipo de cuestiones ofrecen una calificación objetiva (Mendes et al., 2011).

Test final del módulo: Esta autoevaluación es similar a la anterior, se ofrece al alumno una serie de preguntas teóricas del tipo multirrespuesta, pero englobando todo el contenido que se ha estudiado en la asignatura. Por esta razón se habilita una vez desarrollados todos los contenidos, y resueltos los puntos de control y test de las unidades. Los conceptos que aparecen en las preguntas del test final se consideran esenciales para alcanzar los resultados de aprendizaje de la asignatura.

Entrega de ejercicios: En función de los contenidos a estudiar, los estudiantes deberán hacer entrega de ejercicios en algunas unidades. Las asignaturas más teóricas, no tienen este tipo de entregas. En cambio, en las más prácticas, los estudiantes deberán enviar una resolución de los casos prácticos planteados. Tras la corrección de los tutores, el alumno recibirá feedback sobre la tarea enviada. Si el ejercicio no es resuelto correctamente, el estudiante deberá volverlo a enviar siguiendo las instrucciones dadas por el tutor. El caso práctico se calificará cuando el estudiante haya realizado el envío con el ejercicio resuelto correctamente.

Una vez realizadas todas las actividades de la evaluación continua, el estudiante deberá enfrentarse al **examen final** de la asignatura. Durante el curso académico se planifican tres periodos de evaluación para hacer estos exámenes (del Teso et al., 2018). En cada periodo se especifica a cada estudiante los exámenes a los que se pueden presentar, siendo él quien decide si desea presentarse en dicha convocatoria, o si prefiere esperarse a la siguiente. Los alumnos únicamente podrán presentarse al examen final de una asignatura si han finalizado todas las actividades de la evaluación continua y la nota final de ésta es superior a un 5 (sobre un máximo de 10 puntos).

El examen final cuenta con preguntas de diferente tipología (multirrespuesta, resolución numérica, completar huecos...), y con límite de tiempo, el cual varía en función de la asignatura. El objetivo de esta evaluación es similar al del Test Final, se preguntan cuestiones esenciales para garantizar que los alumnos han adquirido los resultados de aprendizaje establecidos. La nota mínima para superar la asignatura debe ser de 5, tanto en la evaluación continua como en el examen final.

3. Metodología

3.1. Descripción de la muestra

La muestra estudiada cuenta con 19 asignaturas distribuidas en 4 títulos de posgrado, y abarca 4 cursos académicos desde 16/17 hasta 19/20. Así, la muestra seleccionada contempla un total de 161 alumnos y 1228 evaluaciones.

Los factores que se estima que pueden tener un impacto en la nota son los siguientes: el título al que pertenecen los alumnos (Máster en Gestión Eficiente del agua urbana, DEU – Diploma de especialización en análisis y diseño de redes de agua, EUE – Experto Universitario en EPANET y EUT – Experto Universitario en Transitorios hidráulicos con Allievi), el periodo de evaluación en el que se realiza el examen final, el tiempo empleado en desarrollar la parte de evaluación continua y el tipo de actividades con las que cuenta cada asignatura (si tiene ejercicio práctico o no).

3.2. Análisis realizado

El análisis estadístico llevado a cabo pretende determinar los posibles factores que influyen en las calificaciones obtenidas por los alumnos y descubrir las posibles causas de las diferencias entre la evaluación continua y el examen final.

Para ello, en primer lugar, se realizará un análisis descriptivo (promedio, desviación estándar...) sobre las notas de la evaluación continua y del examen final. En segundo lugar, tras conocer las características de la muestra, se procederá a analizar si estadísticamente hay diferencias entre las notas medias de los exámenes y la evaluación continua. De este modo, se verificará si efectivamente las diferencias que los tutores creen haber encontrado son estadísticamente significativas. Para ello, se realizará un test de hipótesis, siendo la hipótesis nula que no existen diferencias significativas entre las muestras y la hipótesis alternativa que las medias son diferentes. La prueba se realizará sobre la media (si la muestra sigue una distribución normal) o sobre la mediana en caso contrario.

En el caso en el que efectivamente haya diferencias entre las medias, se llevará a cabo un análisis de varianza (ANOVA) con los factores explicados previamente, con el fin de detectar si impactan en los resultados y tratar de encontrar la razón de dichas diferencias.

Al realizar el ANOVA se deben cumplir los siguientes tres supuestos: (i) las muestras son aleatorias, hecho que se cumple dado que se trata de alumnos distintos; (ii) Las poblaciones son normales; y (iii) las distribuciones deben tener la misma varianza.

Si no se puede garantizar la normalidad de la muestra, se utilizará la prueba de Kruskal-Wallis, que evalúa las diferencias entre las medianas, en lugar de entre las muestras.

El análisis realizado se ha llevado a cabo con el software Statgraphics Centurion XVI.

4. Resultados

En la Tabla 1 se muestran los resultados del análisis descriptivo estadístico realizado sobre las notas de examen (N_EX) y en la evaluación continua (N_CA). Tal y como se puede observar, sí que existe una diferencia notable entre la nota media que los alumnos obtienen en el examen final y la que obtienen en la evaluación continua. De este análisis también se desprende que la variación de notas en la evaluación continua es menor que en los exámenes. En la **¡Error! No se encuentra el origen de la referencia.** se puede observar como la mayoría de las notas de la evaluación continua (parte inferior del gráfico) se sitúan en torno al 9 y el 10 y va decreciendo paulatinamente hasta notas de 7.5. En cambio, en los exámenes finales, aunque hay muchas calificaciones de 10, la muestra se espacia y se presenta un pico en torno a notas en el rango del 7 al 9.

Tabla 1. Descripción y caracterización de la muestra

	<i>Promedio</i>	<i>Desv. Estánd.</i>	<i>Coef. Variación</i>	<i>Mín.</i>	<i>Máx.</i>	<i>Rango</i>	<i>Sesgo Estand.</i>	<i>Curtosis Estand.</i>
N_EX	8,47	1,51	17,85%	1,75	10,0	8,25	-16,24	8,10
N_CA	9,21	0,68	7,426%	5,0	10,0	5,0	-28,86	48,37

Un hecho significativo de la muestra, son los elevados valores de sesgo estandarizado y curtosis que apuntan a que ninguna de las dos muestras sigue una distribución normal.

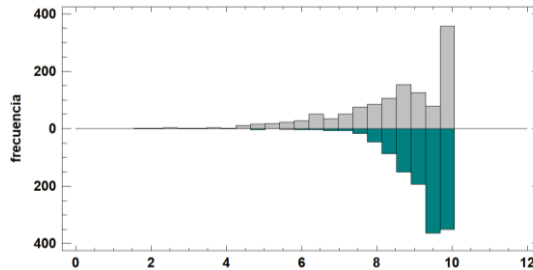


Fig. 1. Gráfico de frecuencias entre la nota del examen final (N_EX), en la parte superior y la de la evaluación continua (N_CA) en la parte inferior

Una vez analizadas las muestras, se realiza el test de hipótesis que determinará si estadísticamente hay diferencia entre los resultados obtenidos en ambos tipos de evaluaciones. Para ello, se realizó un test de medias que detectó que éstas eran significativamente distintas con un nivel de confianza de 95%. Sin embargo, al tratarse de poblaciones no normales y con elevada diferencia en las desviaciones estándar, las hipótesis de partida del test no se cumplen y los resultados no son válidos. Así, el test de hipótesis se realizó sobre las medianas en lugar de las medias (prueba de Kolmogorov-Smirnov). El análisis confirmó que los resultados entre ambas evaluaciones son significativamente distintos con un intervalo de confianza del 95%.

Un análisis en profundidad de la diferencia entre ambas notas, particularizado por asignaturas, indica que, de media, la nota obtenida en el examen final es siempre inferior a la de la evaluación continua con la excepción de dos asignaturas: Modelación avanzada de redes con EPANET y Programación con Toolkit de EPANET, cuya media en el examen es superior a la obtenida tras la evaluación continua. Se estima que el motivo es que se trata de asignaturas complejas que cuentan con casos prácticos muy reales. Se trata de tareas desafiantes que presentan dificultad pero preparan adecuadamente a los alumnos para enfrentarse al examen sin problemas. Por ello, la nota de la evaluación continua es inferior.

Para conocer el efecto de cada factor sobre las muestras, se realizaron análisis ANOVA simples para cada factor. Como las muestras no siguen una distribución estándar, se empleó la prueba de Kruskal-Wallis. La Tabla 2 muestra los resultados obtenidos.

Tabla 2. Resultados ANOVA simple

Factor	Nota Cursosagua	Nota Examen	Nota Final
Titulación	Sí	Sí	Sí
Asignatura	Sí	Sí	Sí
Tiempo de desarrollo	Sí	No	No
Ejercicio	Sí	Sí	No
Convocatoria examen	N/A	Sí	Sí

Sí: el factor afecta significativamente a las muestras. No: el factor no afecta a la muestra. N/A. No aplica. $\alpha=0,05$

No se ha considerado el efecto de la convocatoria de examen sobre la nota de la evaluación continua, ya que hasta que no se ha completado la asignatura no se realiza el examen. Es por ello que dicho factor no influye en esta evaluación.

Posteriormente, en aquellos factores que impactan en las notas, se ha realizado un análisis de las medias entre grupos, para conocer el impacto del factor en la muestra.

Tras analizar los resultados presentados en la Tabla 2 se puede concluir que todos los factores que se había supuesto que afectaban a los resultados, influyen en éstos. La única excepción es el **tiempo de desarrollo de la asignatura**, que únicamente es relevante en la evaluación continua. Se ha comprobado que los estudiantes que más tiempo dedican a revisar el temario obtienen mejores resultados en la evaluación continua. Sin embargo, este hecho no influye en la nota del examen.

En las **titulaciones**, estadísticamente, los alumnos del título Experto en transitorios obtienen ligeramente peores resultados (8.5 de media) con respecto a los del DEU (9.1). Los alumnos del EUE y del Máster obtienen resultados intermedios. Estas diferencias se mantienen en la evaluación continua, los resultados de los exámenes y, por tanto, en la nota final. El motivo parece apuntar a que el Experto en transitorios trata contenidos muy específicos en los que los alumnos no tienen conceptos previos. Asimismo, las asignaturas son eminentemente prácticas con abundantes ejercicios y casos prácticos, fomentando la adquisición de los resultados de aprendizaje.

Por **asignaturas**, Indicadores de Gestión y Resolución de casos prácticos de transitorios hidráulicos con Allievi son, con diferencia, las que tienen las notas finales más bajas. Por el contrario, Gestión de la demanda y Análisis de redes de saneamiento con SWMM son las que mejores calificaciones tienen.

Al analizar la **diferencia de notas entre las dos evaluaciones**, Evaluación y control de pérdidas es la asignatura con mayor diferencia media entre las notas de ambos tipos de evaluaciones, obteniendo de media en el examen 2.3 puntos menos. Seguida por Hidráulica básica en lámina libre e hidrología urbana e Indicadores de gestión, ambas con una media en el examen de 1.80 puntos por debajo de la puntuación de la evaluación continua.

Las asignaturas de modelación relacionadas con los softwares de EPANET y SWMM son las que menores diferencia presentan entre los dos tipos de evaluación, con la excepción del módulo de Gestión Patrimonial de infraestructuras que es el que menor diferencia presenta.

Por **convocatoria**, los resultados de la primera y tercera convocatoria son similares. Sin embargo, la segunda convocatoria obtiene mejores calificaciones. Probablemente, esto se deba a que los alumnos ya conocen el sistema de exámenes y se preparan mejor para éstos, mientras que en la primera convocatoria no. La tercera convocatoria, al ser la última, los alumnos (en general) suelen llegar con más prisa a las últimas asignaturas, presentándose a un mayor número de exámenes finales. Este puede ser el motivo por el que los resultados son ligeramente más bajos que en la segunda convocatoria.

Finalmente, se ha evaluado también la **influencia de los ejercicios**, tanto en las notas medias de la evaluación continua como en el examen final. Los resultados demuestran que las asignaturas sin ejercicios prácticos obtienen de media 0.2 puntos más en la evaluación continua. Este hecho puede ser explicado por los distintos pesos que tienen las áreas de la evaluación continua en función de si se cuenta con entrega de ejercicios o no.

Sin embargo, en los exámenes finales, aquellas asignaturas que cuentan con entrega de ejercicios a lo largo de la evaluación continua obtienen de media notas superiores en 0.3 puntos. Este hecho indica que la entrega de ejercicios prepara mejor a los estudiantes para enfrentarse a los exámenes.

5. Conclusiones

Tras el análisis estadístico realizado, se concluye, en primer lugar, que la inclusión de ejercicios prácticos en la evaluación continua prepara mejor a los alumnos para alcanzar los objetivos de aprendizaje. Es por ello que las futuras mejoras deben estar orientadas a incluir ejercicios prácticos en todas las asignaturas con el fin de fijar los conocimientos adquiridos.

En segundo lugar, es también relevante el hecho de que los exámenes de la segunda convocatoria obtienen mejores resultados. Así, se concluye que se debe incidir más en el seguimiento del alumnado para motivarles a realizar un estudio continuado, y equilibrar el número de exámenes finales realizados entre el segundo y el tercer periodo de evaluación, evitando la acumulación de asignaturas al final del curso académico.

Finalmente, para que los resultados de la evaluación continua y el examen final no sean tan diferentes, se plantea modificar los test de las unidades y el test final para que contengan menos preguntas teóricas y sean más aplicados. Un factor a tener en cuenta, pero que no se ha podido cuantificar en el estudio estadístico, es el hecho de que el examen final tiene un

tiempo límite de realización, mientras que las actividades de la evaluación continua no. Los futuros desarrollos van a plantearse para introducir también el factor tiempo en los test de las unidades y el test final. De este modo, con ambas modificaciones el alumnado estará mejor preparado para enfrentarse al examen final, que cuenta con un carácter más práctico y con límite temporal.

Referencias

- del Teso, R., Estruch, E., Gómez, E., & Soriano, J. (2018). Sistema de evaluación para la formación a distancia de profesionales. In *IN-RED 2018*. Valencia.
- García-Beltrán, Á., Martínez, R., Jaén, J.-A., & Tapia, S. (2006). La autoevaluación como actividad docente en entornos virtuales de aprendizaje/enseñanza. *RED. Revista de Educación a Distancia*. <https://doi.org/10.6018/red/50/14>
- Gibbs, G., & Simpson, C. (2009). *Condiciones para una evaluación continuada favorecedora del aprendizaje*. (Octaedro, Ed.). Barcelona: Cuadernos de docencia universitaria.
- Rodríguez, G., & Ibarra, M. S. (2011). *E-evaluación orientada al e-aprendizaje estratégico en Educación Superior*. (Narcea, Ed.). Madrid.
- VICERRECTORADO DE CALIDAD. UNIVERSIDAD DE EXTREMADURA. "Consejos para la creación y gestión de MOOC - Miriada X". <https://www.unex.es/organizacion/gobierno/vicerrektorados/vicecal/archivos/ficheros/moocs/Consejos_rapidos_para_formato_MOOC.pdf/view> [Consulta: 20 de mayo de 2018] [Ángel González de la Fuente y David Carabantes Alarcón]